Metadata & Findability

Author: Bruna Piereck and Alexander Botzki

Version: 092024

CC BY NC SA

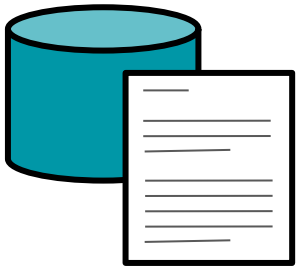
Findability

Human findable

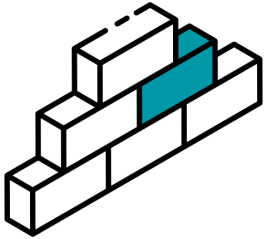
Machine findable

What do you know about metadata?

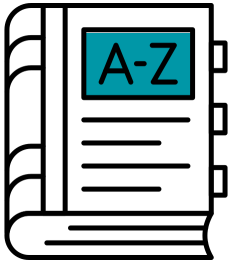
What is Metadata



Data to describe Data



Structured data

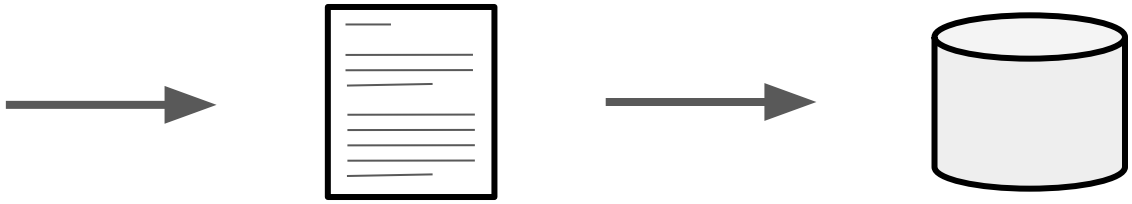


Descriptor with clear meaning

“Love note to the future about data”

General structure of Metadata

Property A
Property B
Property C
Property D
etc



Properties

Profile

Type

One example .... →

Type

Profile

Recipe

A Schema.org Type

Thing > CreativeWork > HowTo > Recipe

[more...]

A recipe. For dietary restrictions covered by the recipe, a few common restrictions are enumerated via suitableForDiet. The keywords property can also be used to add more detail.

Property	Expected Type	Description
Properties from Recipe		
cookTime	Duration	The time it takes to actually cook the dish, in ISO 8601 duration format.
cookingMethod	Text	The method of cooking, such as Frying, Steaming, ...
nutrition	NutritionInformation	Nutrition information about the recipe or menu item.
recipeCategory	Text	The category of the recipe—for example, appetizer, entree, etc.
recipeCuisine	Text	The cuisine of the recipe (for example, French or Ethiopian).
recipeIngredient	Text	A single ingredient used in the recipe, e.g. sugar, flour or garlic. Supersedes ingredients.
recipeInstructions	CreativeWork or ItemList or Text	A step in making the recipe, in the form of a single item (document, video, etc.) or an ordered list with HowToStep and/or HowToSection items.
recipeYield	QuantitativeValue or Text	The quantity produced by the recipe (for example, number of people served, number of servings, etc).
suitableForDiet	RestrictedDiet	Indicates a dietary restriction or guideline for which this recipe or menu item is suitable, e.g. diabetic, halal etc.
Properties from HowTo		
estimatedCost	MonetaryAmount or Text	The estimated cost of the supply or supplies consumed when performing instructions.
performTime	Duration	The length of time it takes to perform instructions or a direction (not including time to prepare the supplies), in ISO 8601 duration format.
prepTime	Duration	The length of time it takes to prepare the items to be used in instructions or a direction, in ISO 8601 duration format.
	CreativeWork or HowToSection or	A single step item (as HowToStep, text, document, video, etc.) or a HowToSection. Supersedes steps.

Type

Profile

Recipe

A Schema.org Type

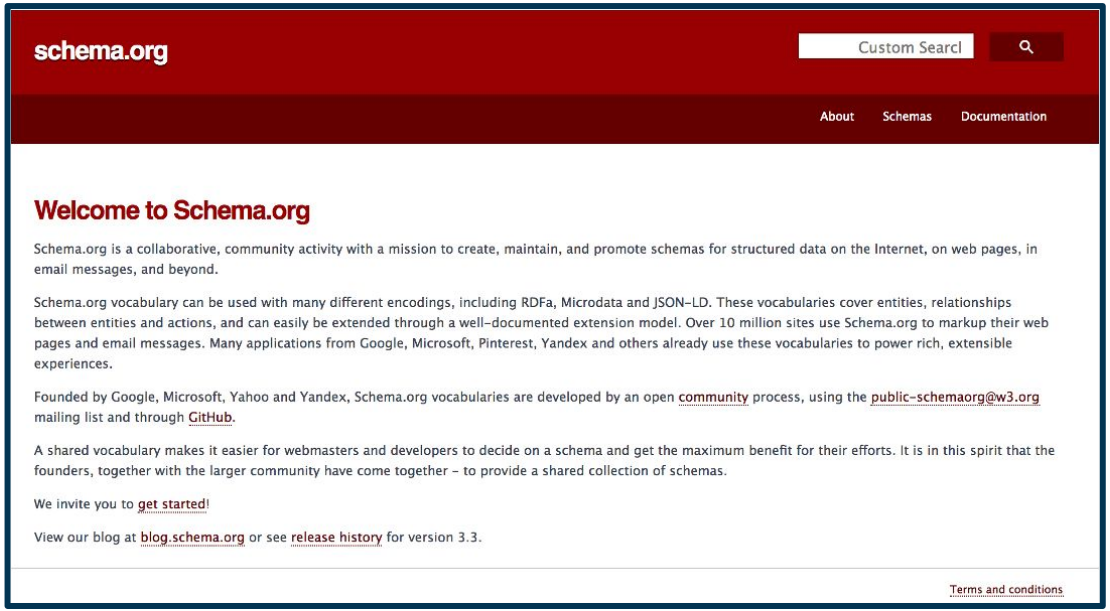
Thing > CreativeWork > HowTo > Recipe

[more...]

A recipe. For dietary restrictions covered by the recipe, a few common restrictions are enumerated via suitableForDiet. The keywords property can also be used to add more detail.

Property	Expected Type	Description
Properties from Recipe		
cookTime	Duration	The time it takes to actually cook the dish, in ISO 8601 duration format.
cookingMethod	Text	The method of cooking, such as Frying, Steaming, ...
nutrition	NutritionInformation	Nutrition information about the recipe or menu item.
recipeCategory	Text	The category of the recipe—for example, appetizer, entree, etc.
recipeCuisine	Text	The cuisine of the recipe (for example, French or Ethiopian).
recipeIngredient	Text	A single ingredient used in the recipe, e.g. sugar, flour or garlic. Supersedes ingredients.
recipeInstructions	CreativeWork or ItemList or Text	A step in making the recipe, in the form of a single item (document, video, etc.) or an ordered list with HowToStep and/or HowToSection items.
recipeYield	QuantitativeValue or Text	The quantity produced by the recipe (for example, number of people served, number of servings, etc).
suitableForDiet	RestrictedDiet	Indicates a dietary restriction or guideline for which this recipe or menu item is suitable, e.g. diabetic, halal etc.
Properties from HowTo		
estimatedCost	MonetaryAmount or Text	The estimated cost of the supply or supplies consumed when performing instructions.
performTime	Duration	The length of time it takes to perform instructions or a direction (not including time to prepare the supplies), in ISO 8601 duration format.
prepTime	Duration	The length of time it takes to prepare the items to be used in instructions or a direction, in ISO 8601 duration format.
	CreativeWork or HowToSection or	A single step item (as HowToStep, text, document, video, etc.) or a HowToSection. Supersedes steps.

Schema.org



- A consortium:
  - Describe online information
  - Improve search efficacy
- Controlled vocabulary
  - Metadata for the world
  - Several profiles

Schema.org

Types:

Close hierarchy / Open hierarchy

- Thing -
  - Action +
  - BioChemEntity +
  - CreativeWork +
  - Event +
  - Intangible +
  - MedicalEntity +
  - Organization +
  - Person +
  - Place +
  - Product +
  - Taxon

Event + → More specific Types

- BusinessEvent
- ChildrensEvent
- ComedyEvent

etc …

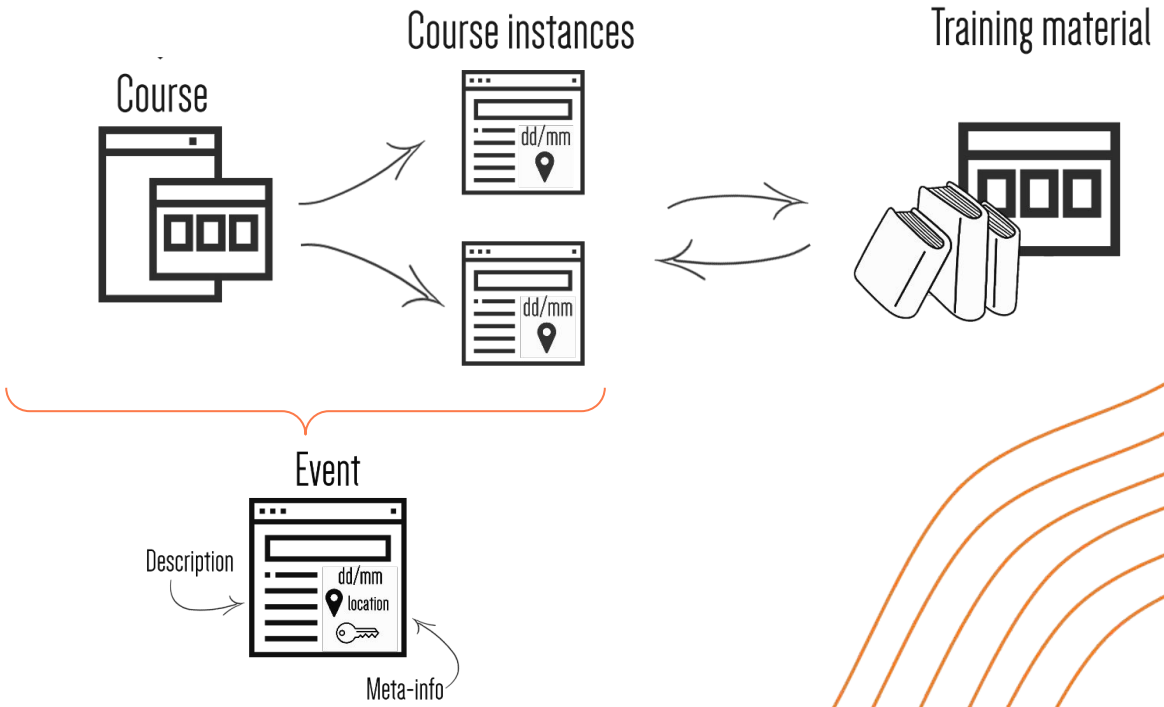
BusinessEvent →

Profile

Property	Expected type	Description
about	thing	…
actor	person	…
attendee	Person or Org	…

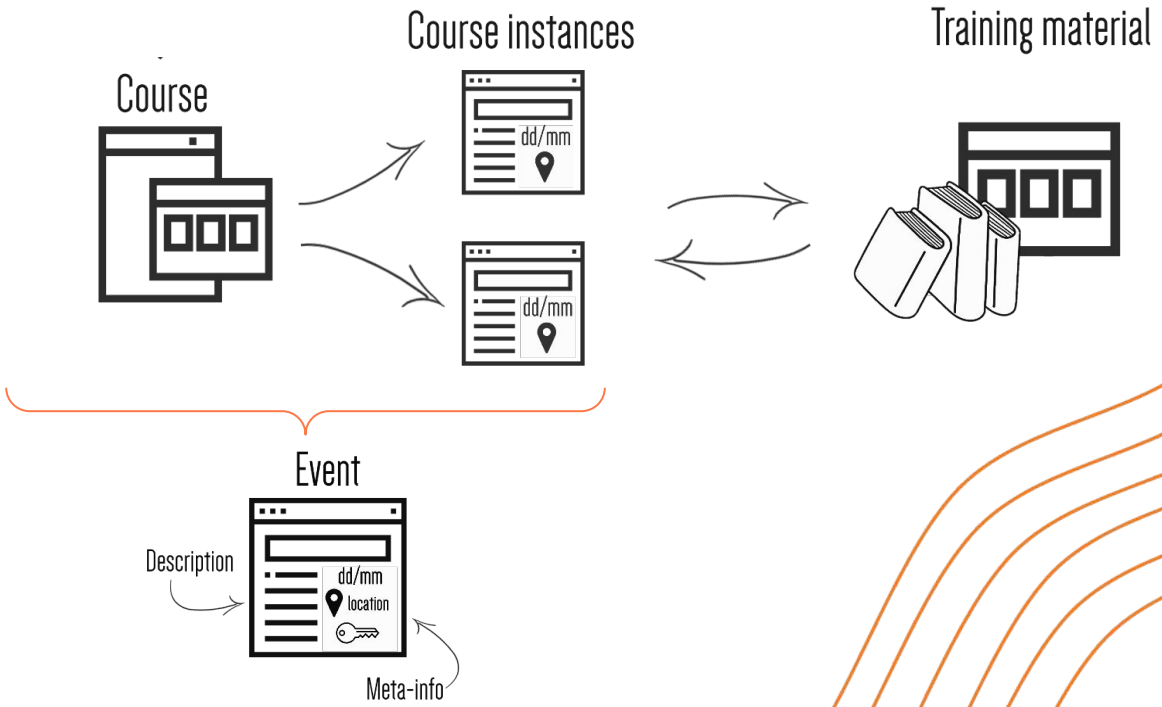
What training types exist

Types for training



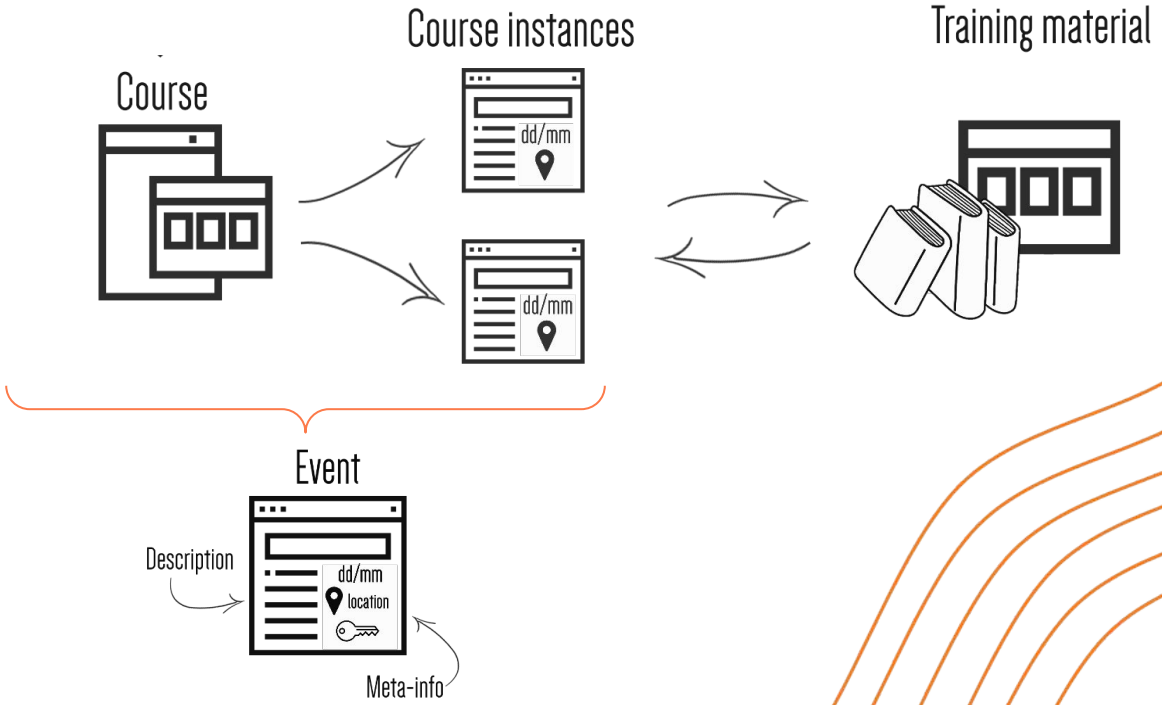
Types for training

- Course:
  - FAIR training material made *by design*

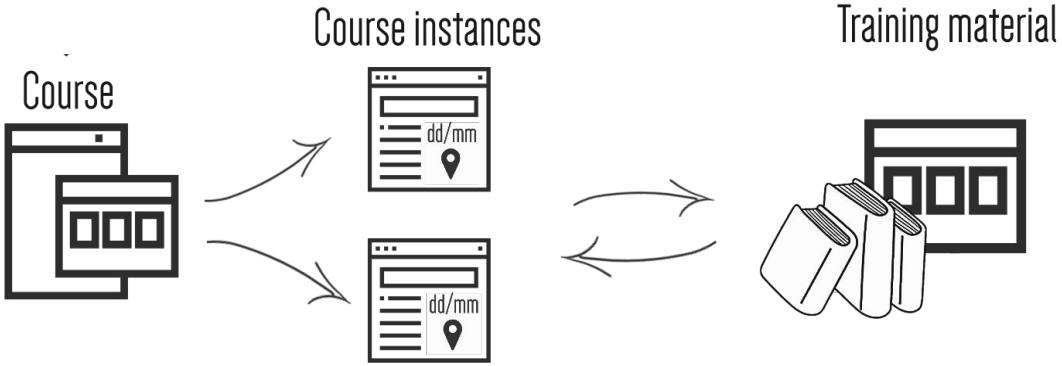


Types for training

- Course:
  - FAIR training material made *by design*
- Course instance:
  - Sept 2024 - SE
  - Sept 2025 - BE
  - Sept 2026 - IT



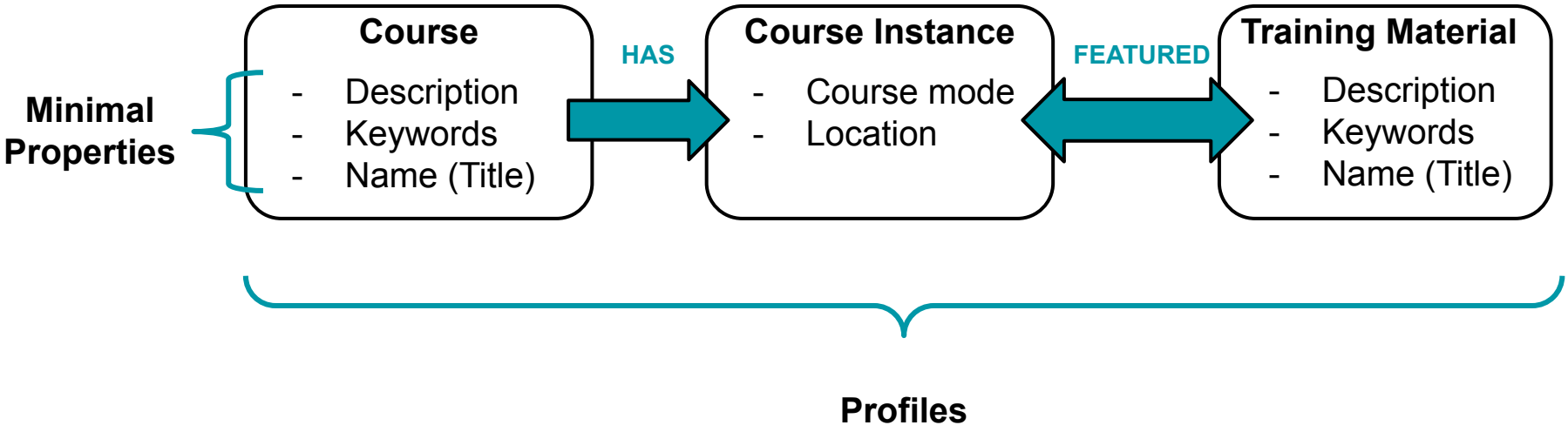
What properties there are in each profile



?

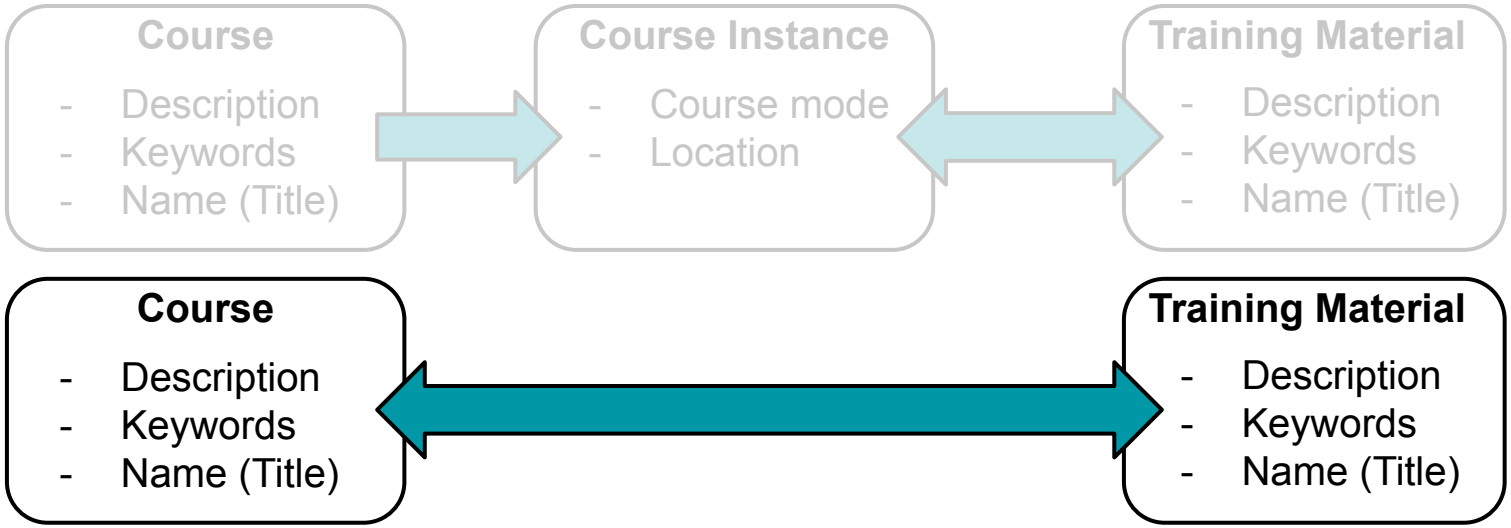
example from Niall Beard
https://bioschemas.org/profiles/

Training minimum properties



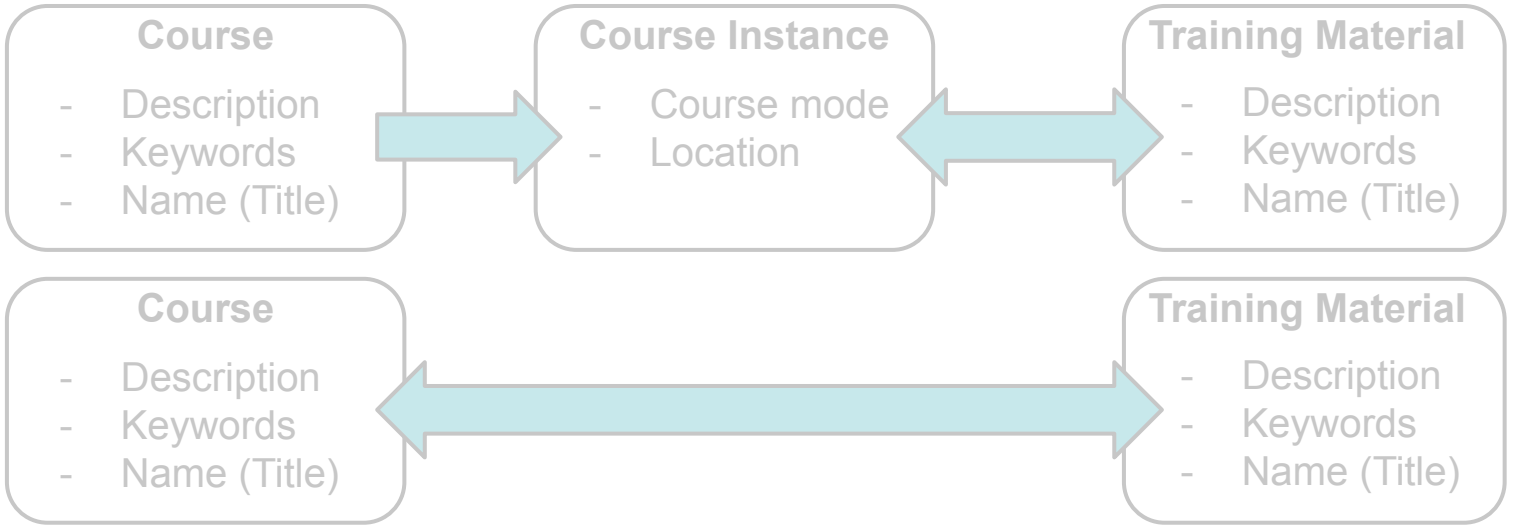
example from Niall Beard
https://bioschemas.org/profiles/

Training minimum properties



example from Niall Beard
https://bioschemas.org/profiles/

Training minimum properties



Training Material

- Description
- Keywords
- Name (Title)

Find out all the properties

Bioschemas

https://bioschemas.org/profiles/

Home | Getting Started | Specifications | Deploy & Develop | Community | About

Schema.org

Bioschema .org

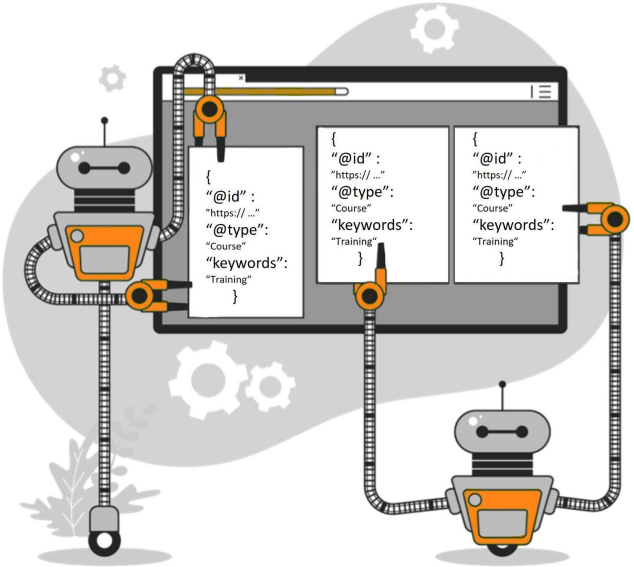
Bioschemas Profiles

Propose a new profile

The Bioschemas' profiles define a community agreed layer over the Schema.org vocabulary, providing additional constraints. These constraints capture (i) the information properties agreed by the community which are minimum (M), recommended (R), or optional (O), (ii) the cardinality of the property, i.e. whether it is expected to occur once or many times, and (iii) associated controlled vocabulary terms drawn from existing ontologies.

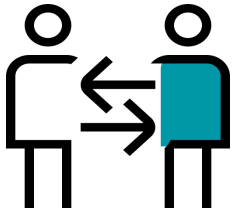
These guidelines provide community best practice that lead to consistent markup of life sciences resources on the Web. Comments and discussions on the mailing list or GitHub issue tracker are encouraged.

Impacts of Metadata

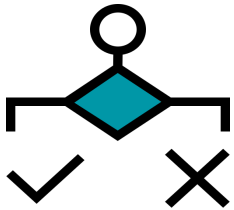


Metadata impact

Findability: Findable by search engines

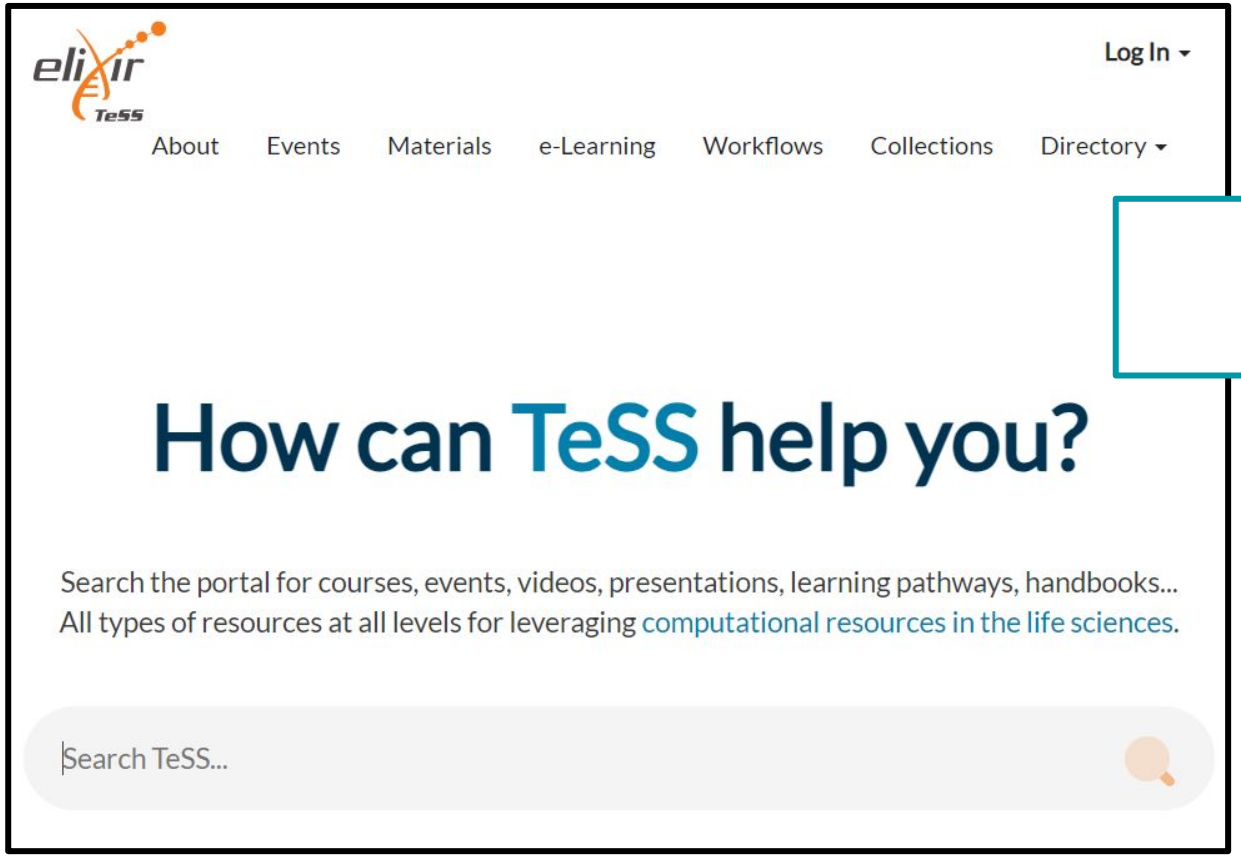


Sharing: Structured information to evaluate



Decision making: Targeted judgment
Evaluation on reusability

Training e-Support System



Aggregator

Metadata keeper !!!

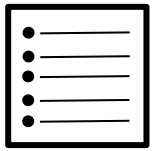
What you can register and find is TeSS portal?

Adapted from @Sketchepedia - freepick.com

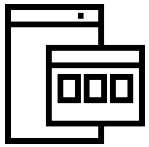
What you can register and find is TeSS portal?

- Training material

Course plan

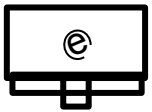


Tutorial



Course content

Collections

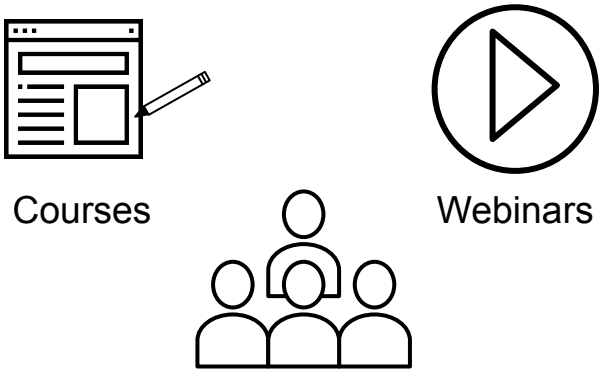


e-Learning course

Adapted from @Sketchepedia - freepick.com

What you can register and find is TeSS portal?

- Training material
- Events



Adapted from @Sketchepedia - freepick.com

Who is feeding TeSS portal?

Adapted from @Sketchepedia - freepick.com

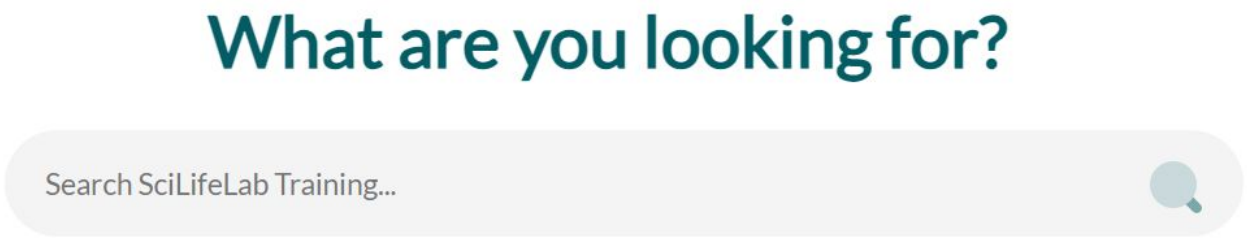
Who is feeding TeSS portal?

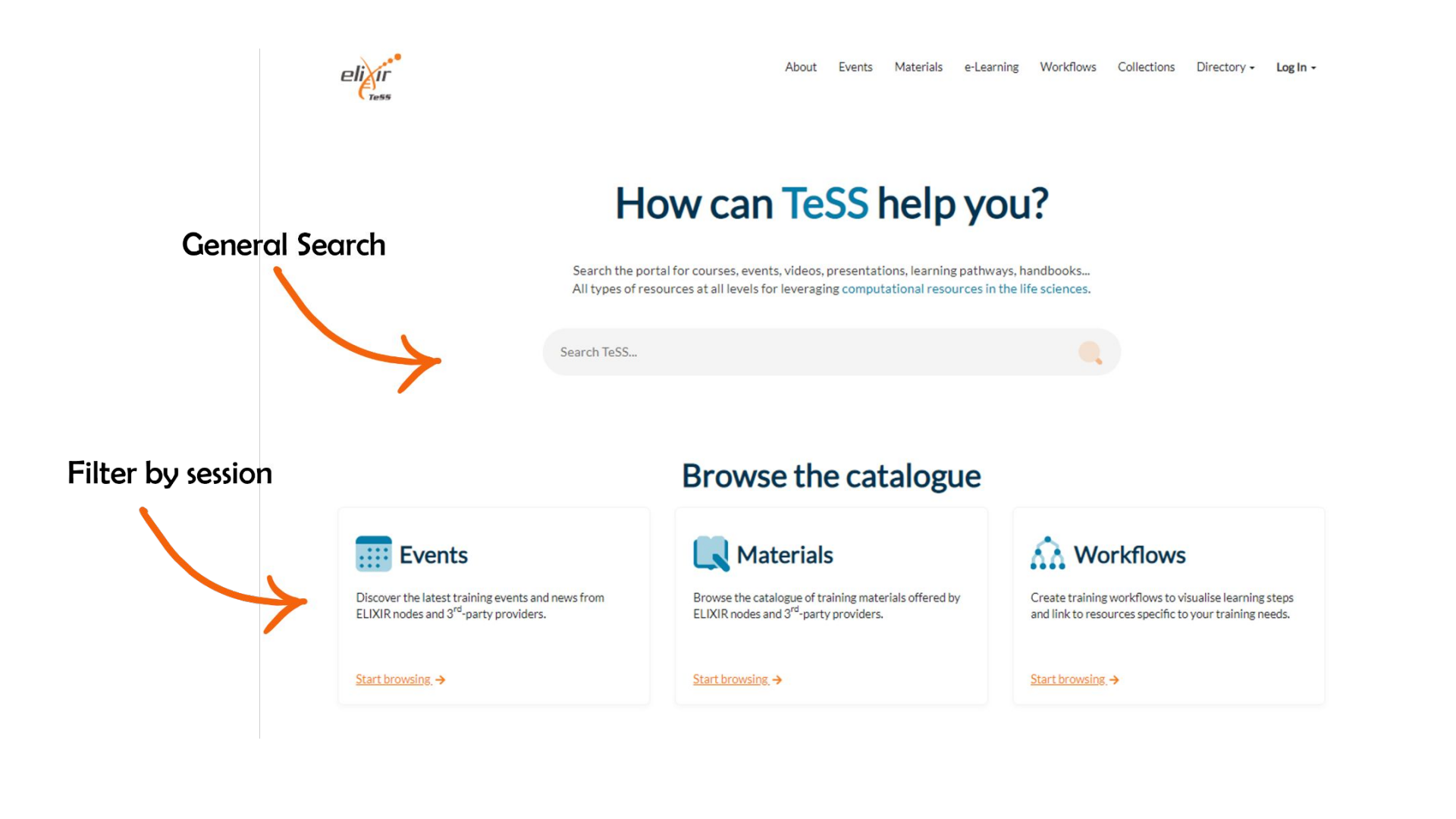
- 122 Content providers
  - Academic institutions
  - non-profit organisations
  - Other portals

- Registered people

Adapted from @Sketchepedia - freepick.com

TeSS is open source

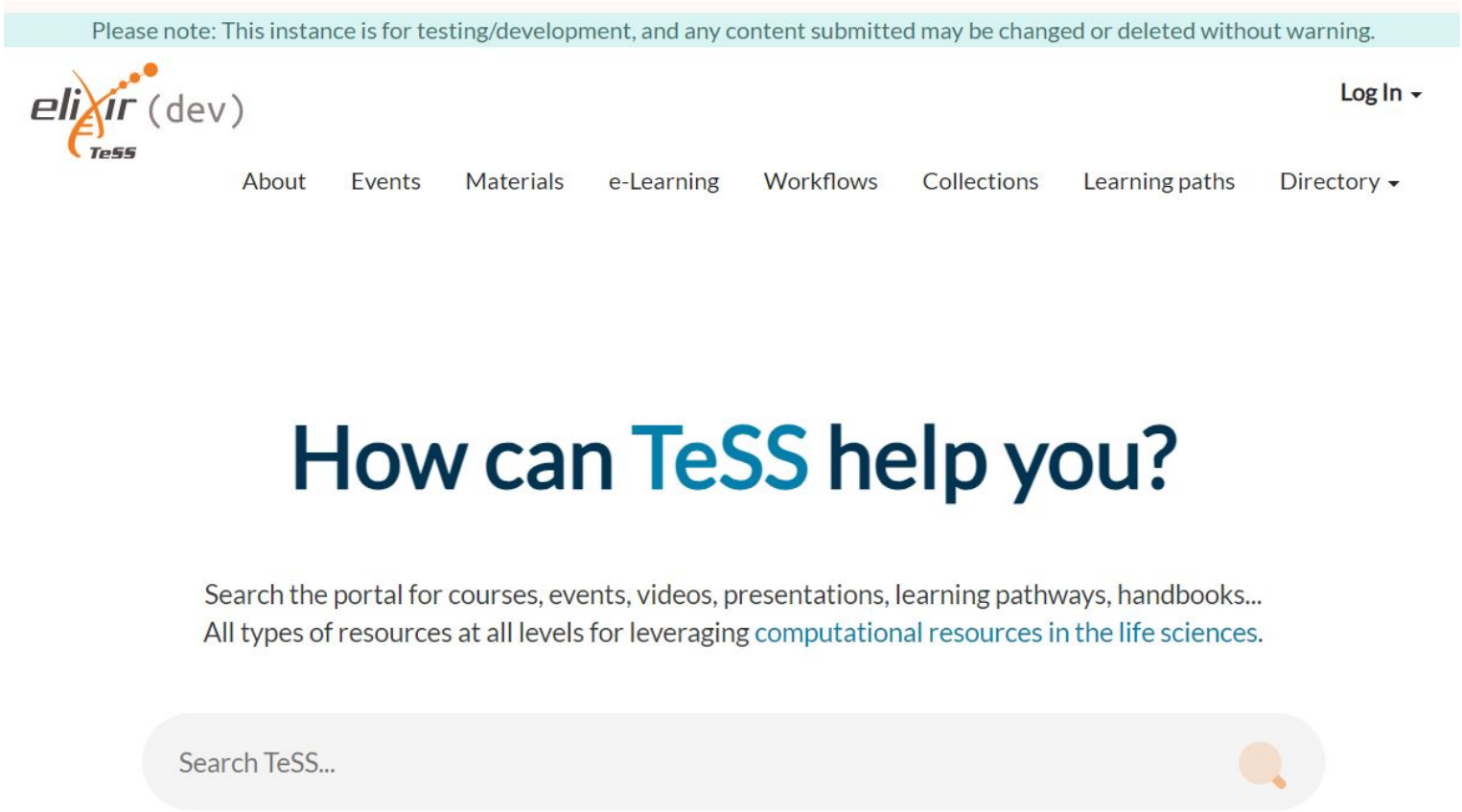



About
Events
Materials
e-Learning
Workflows
Collections
Directory
Log In
How can TeSS help you?
Search the portal for courses, events, videos, presentations, learning pathways, handbooks...
All types of resources at all levels for leveraging computational resources in the life sciences.
Search TeSS...
General Search
Filter by session
Browse the catalogue
Events
Discover the latest training events and news from ELIXIR nodes and 3rd-party providers.
Start browsing
Materials
Browse the catalogue of training materials offered by ELIXIR nodes and 3rd-party providers.
Start browsing
Workflows
Create training workflows to visualise learning steps and link to resources specific to your training needs.
Start browsing

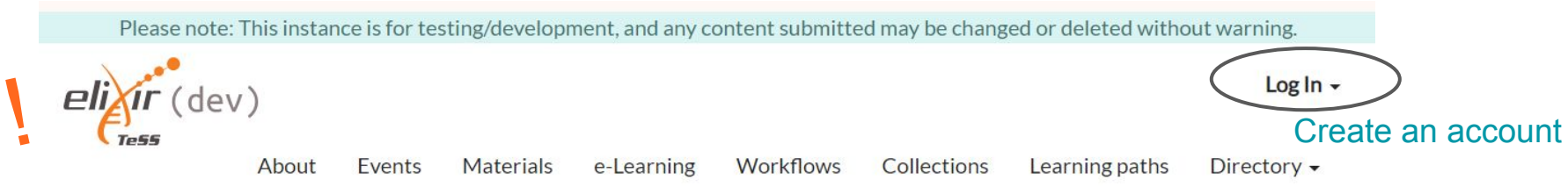
Activity:

1. Find a training material of intermediate difficulty for single cell data analysis
2. Find the event “FAIR training material made *by design*” course
3. Find active training material of NBIS
4. List how many archived materials NBIS has
5. Find when and where VIB will offer a super nice AI event

Let’s prepare to register our own



Let's prepare to register our own



!

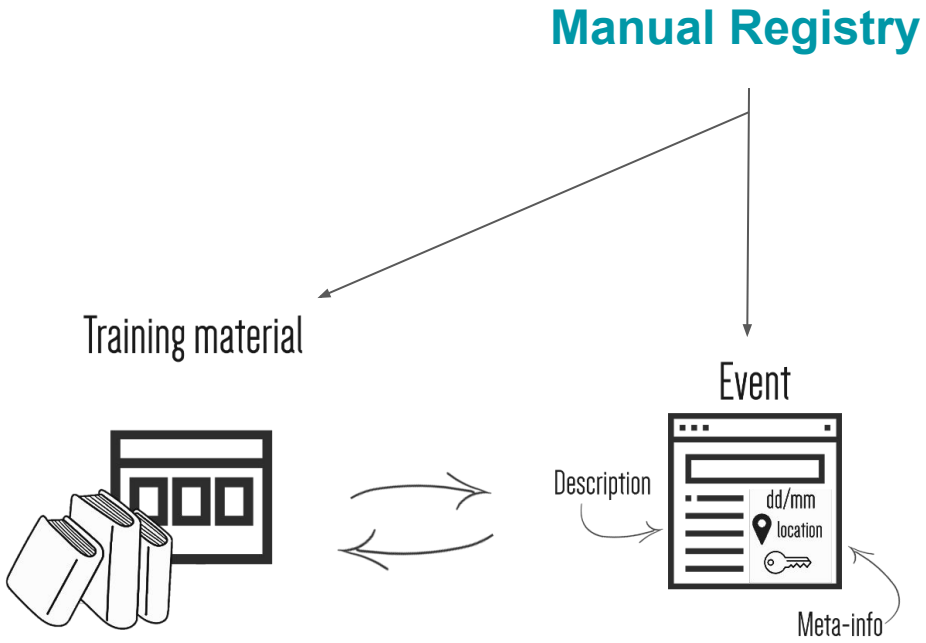
https://dev.tess.elixir-europe.org/

How can TeSS help you?

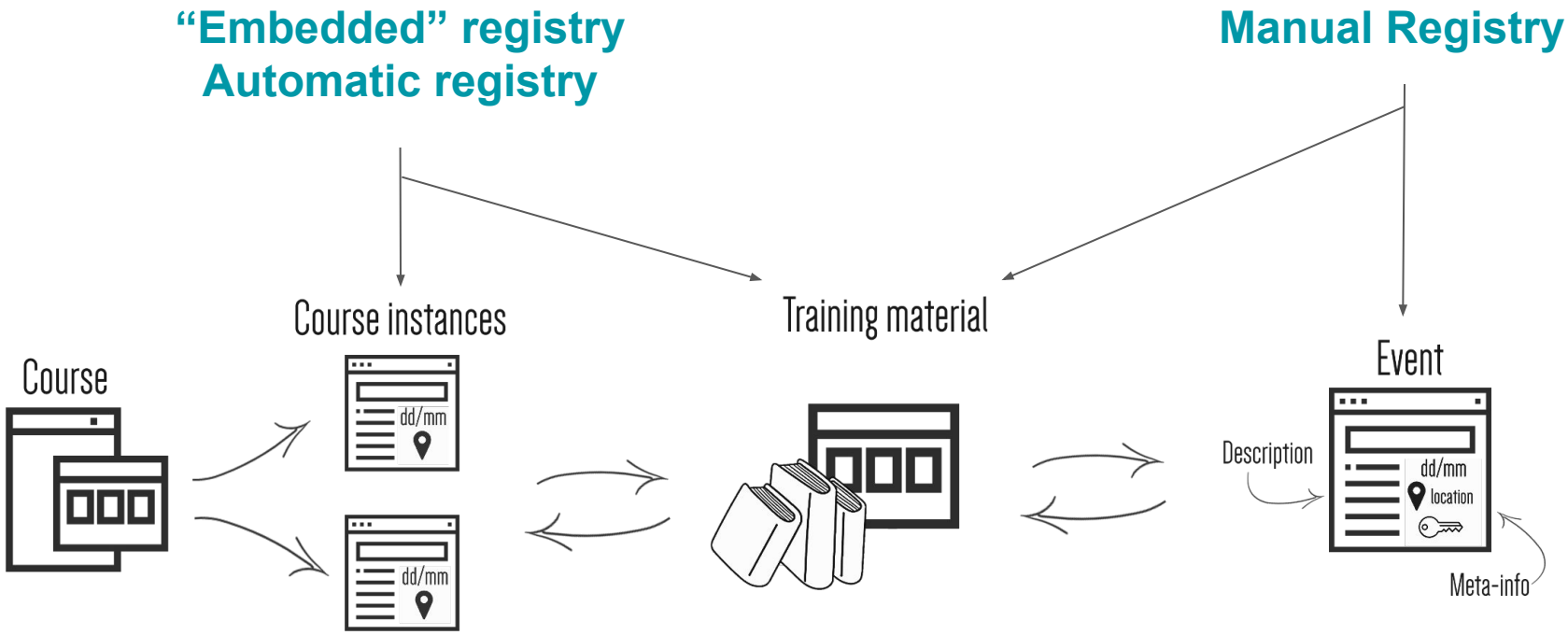
Search the portal for courses, events, videos, presentations, learning pathways, handbooks...
All types of resources at all levels for leveraging computational resources in the life sciences.

Search TeSS...

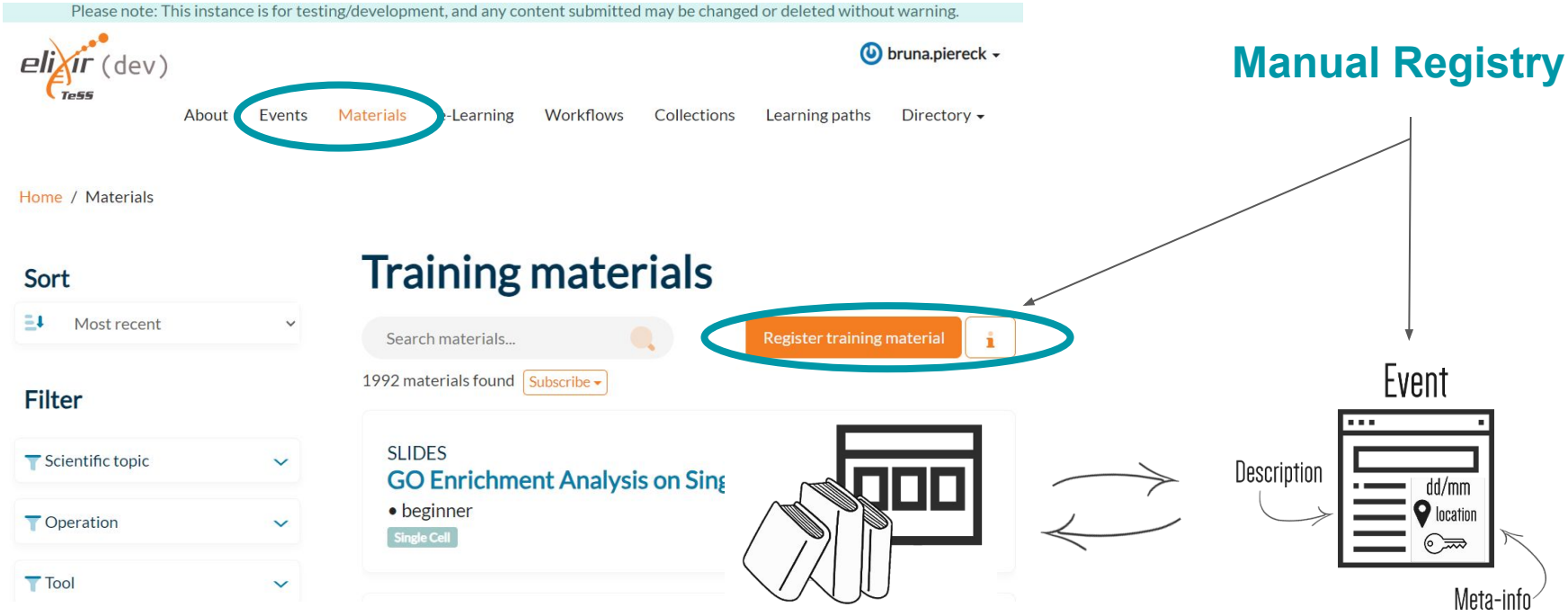
Many ways to register



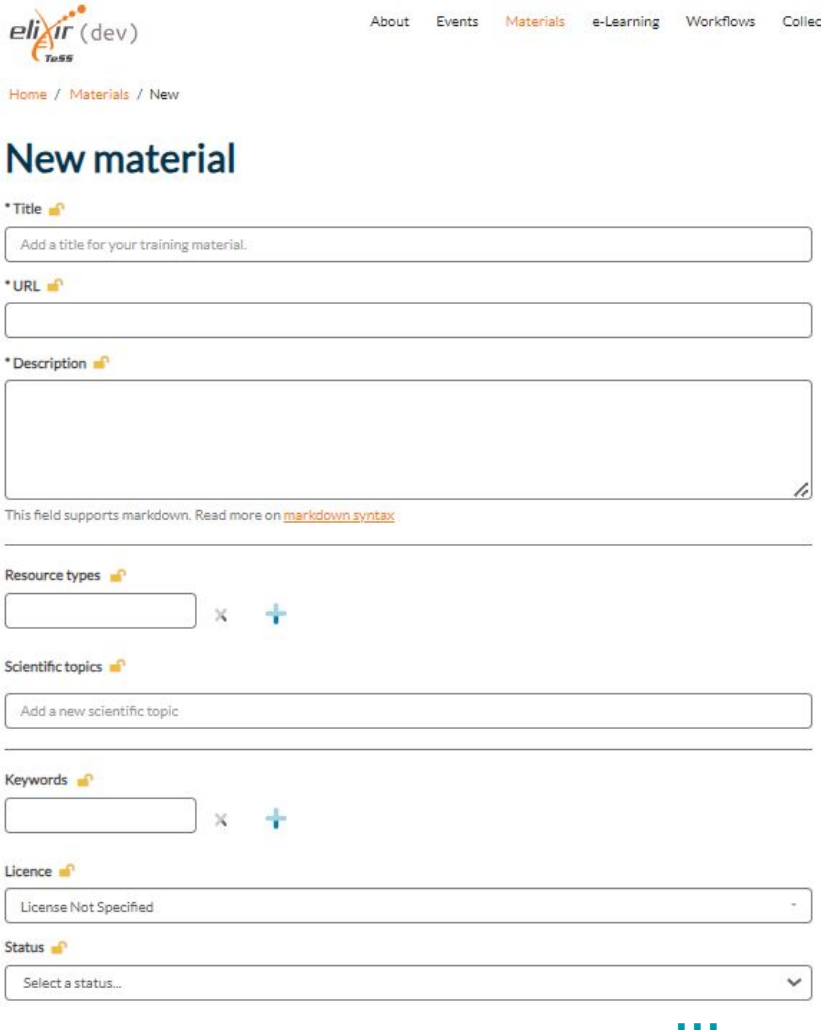
Many ways to register



Manual Registration



Structured data in disguise

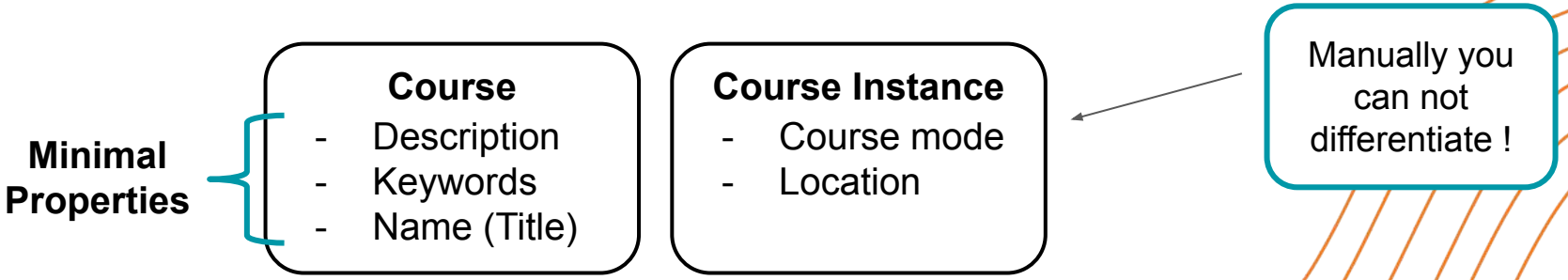


Properties

...

Activity:

1. Create an imaginary event on dev.tess.elixir-europe.org, later we will link our training material to this event.
   a. Annotate as much metadata as you can for your training
   b. Try to think of what usually you would annotate
   c. I trust you can do more than the minimum



Structured data

JSON object

- Data is in key/value pairs
  - KEY: VALUE

"@type": "Person"
"name": "Bruna Piereck"

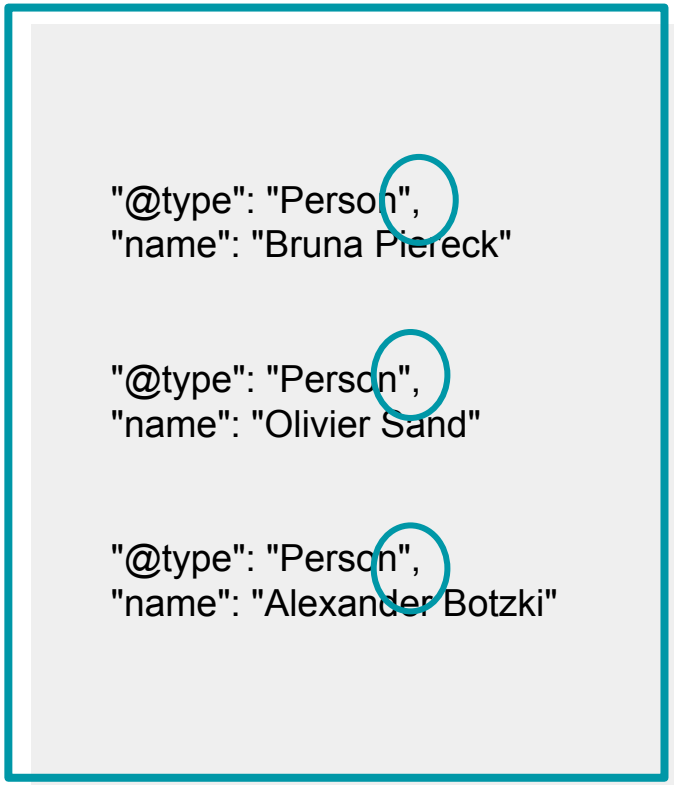
"@type": "Person"
"name": "Olivier Sand"

"@type": "Person"
"name": "Alexander Botzki"

Structured data

JSON object

- Data is in key/value pairs
- Data is separated by commas



Structured data

JSON object

- Data is in name/value pairs
- Data is separated by commas
- Square brackets hold arrays (collection of elements)
    - Name : [collection]

```
"author": [

    "@type": "Person",
    "name": "Bruna Piereck"


    "@type": "Person",
    "name": "Olivier Sand"


    "@type": "Person",
    "name": "Alexander Botzki"

],
```

Structured data

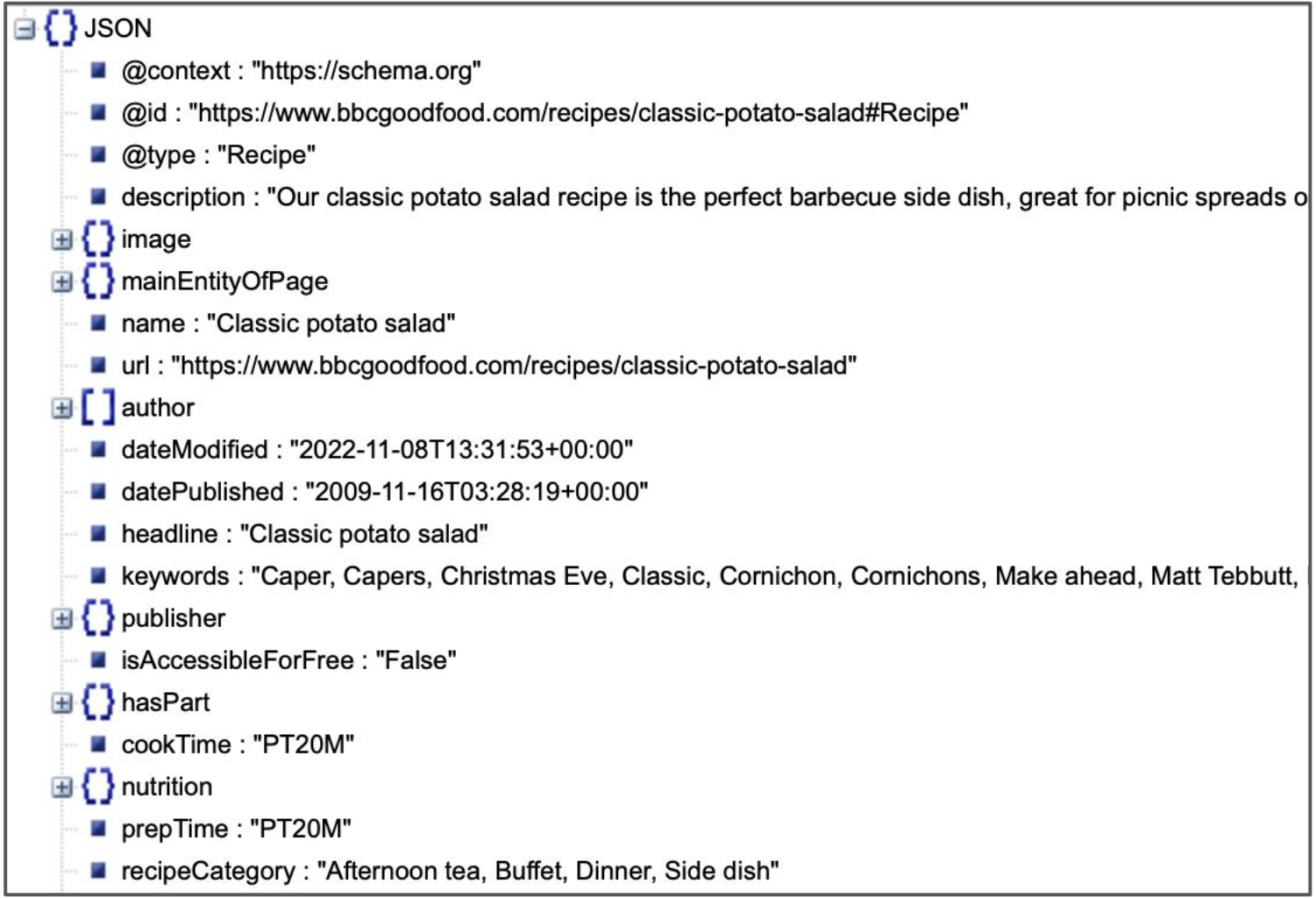
JSON object

- Data is in key/value pairs
- Data is separated by commas
- Square brackets hold arrays (collection of elements)
- Curly braces hold objects

```
"author": [
   {
     "@type": "Person",
     "name": "Bruna Piereck"
   },
   {
     "@type": "Person",
     "name": "Olivier Sand"
   },
   {
     "@type": "Person",
     "name": "Alexander Botzki"
   }
 ],
```

Example of structured data

Example of JSON-LD



Example of structured data

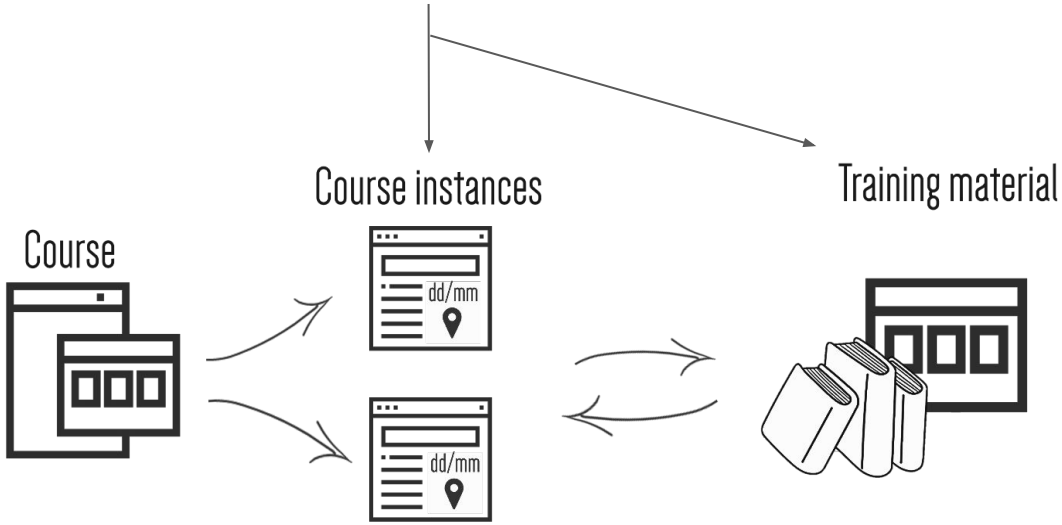
Example of JSON-LD

- JSON
  - @context : "https://schema.org"
  - @id : "https://www.bbcgoodfood.com/recipes/classic-potato-salad#Recipe"
  - @type : "Recipe"
  - description : "Our classic potato salad recipe is the perfect barbecue side dish, great for picnic spreads o
  - image
  - mainEntityOfPage
  - name : "Classic potato salad"
  - url : "https://www.bbcgoodfood.com/recipes/classic-potato-salad"
  - author
  - dateModified : "2022-11-08T13:31:53+00:00"
  - datePublished : "2009-11-16T03:28:19+00:00"
  - headline : "Classic potato salad"
  - keywords : "Caper, Capers, Christmas Eve, Classic, Cornichon, Cornicl
  - publisher
  - isAccessibleForFree : "False"
  - hasPart
  - cookTime : "PT20M"
  - nutrition
  - prepTime : "PT20M"
  - recipeCategory : "Afternoon tea, Buffet, Dinner, Side dish"

Adapted from @Sketchepedia - freepick.com

“Embedded” Registration

“Embedded” registry
Automatic registry



- JSON object -

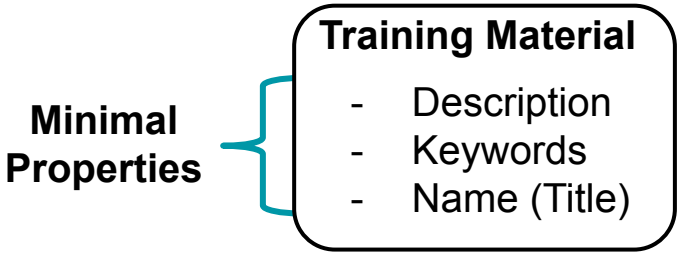
{}JSON
Key:descriptor
Key:descriptor
.
.
.

```
{
  "@context": "https://schema.org/",
  "@type": "LearningResource",
  "@id": "https://elixir-europe-training.github.io/ELIXIR-TrP-TeSS/",
  "http://purl.org/dc/terms/conformsTo": {
    "@type": "CreativeWork",
    "@id": "https://bioschemas.org/profiles/TrainingMaterial/1.0-RELEASE"
  },
  "description": "TeSS, how can I help you? This is our interactive hands-on course about efficient use
  "keywords": "FAIR, OPEN, Bioinformatics, Teaching, TeSS",
  "name": "TeSS, how can I help you?",
  "license": "https://creativecommons.org/licenses/by/4.0/",
  "educationalLevel": "beginner",
  "competencyRequired": "none",
  "teaches": [
    "search events and material in TeSS via direct and faceted search",
    "add manually and automatically events and material to TeSS",
    "extract events and material from TeSS by using TeSS widgets"
  ],
  "audience": "training providers",
  "inLanguage": "en-US",
  "learningResourceType": [
    "tutorial"
  ],
  "author": [
    {
      "@type": "Person",
      "name": "Bruna Piereck"
    },
    {
      "@type": "Person",
      "name": "Olivier Sand"
```

GitHub Octocat, by simon oxley, attribution and non-derivative license

Activity:

1. Annotate your training material metadata
    a. Annotate as much metadata as you can for your training material
    b. Try to think of what usually you would annotate
    c. I trust you can do more than the minimum
    d. Link this material to the event



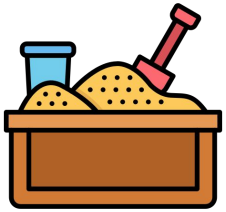
Activity:

1. Annotate your training material metadata
    a. Annotate as much metadata as you can for your training material
    b. Try to think of what usually you would annotate
    c. I trust you can do more than the minimum
    d. Link this material to the event
        i. Manual: the tab is a look-up type the title of your material to embed it in the event.

If you are using GitHub, use the template to include all the metadata

GitHub Octocat, by simon oxley, attribution and non-derivative license

Activity:



@Maan Icons in freepik

1. Extract your JSON file from TeSS
    a. Validate the individual page with the schema.org validator by pasting the URL into the Fetch URL tab.
    b. You will get the the HTML code of the web site on the left side and a list of type from schema.org.
    c. In the HTML code, look for a script element with the attribute type=”application/ld+json”. This script element could be a child of the head element of the HTML page. This results in the following script HTML element.
    d. Copy the JSON object and paste it into a file called course-descr.json

GitHub Octocat, by simon oxley, attribution and non-derivative license

Demo:

Scrape the metadata about the material into TeSS

We will demo how you could set up an automated mechanism to get training material into TeSS.